AL-TR-1991-0059

AD-A241 996

OBSERVING SYSTEM FOR CRITIQUE, ADVICE AND REVIEW (OSCAR)

ARMSTRONG LABORATORY

Richard W. Obermayer

VRC Corporation
68 Long Court, Suite 2A
Thousand Oaks, CA 91360-6084

HUMAN RESOURCES DIRECTORATE
AIRCREW TRAINING RESEARCH DIVISION
Williams Air Force Base, AZ 85240-6457

September 1991

Final Technical Report for Period May 1988 - April 1991

Approved for public release; distribution is unlimited.

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5000

NOTICES

This report has been reviewed and is approved for publication.

RICHARD A. THURMAN
Contract Monitor

DEE H. ANDREWS, Technical Director
Aircrew Training Research Division

JOHN H. FULLER, JR., Colonel, USAF
Chief, Aircrew Training Research Division

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1991	Final – May 1988 – April 1991

4. TITLE AND SUBTITLE

Observing System for Critique, Advice and Review (OSCAR)

5. FUNDING NUMBERS

C – F33615-88-C-0005
PE – 62205F
PR – 1123
TA – 35
WU – 13

6. AUTHOR(S)

Richard W. Obermayer

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

VRC Corporation
68 Long Court, Suite 2A
Thousand Oaks, CA 91360-6084

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

Armstrong Laboratory
Human Resources Directorate
Aircrew Training Research Division
Williams Air Force Base, AZ 85240-6457

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

AL-TR-1991-0059

11. SUPPLEMENTARY NOTES

Armstrong Laboratory Technical Monitor: Dr. Richard A. Thurman, (602) 988-6561

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The Observing System for Critique, Advice and Review (OSCAR) is a test system to provide expert critique of performance as training feedback. OSCAR was developed in three phases. The Phase I effort provided initial development and a demonstration of the feasibility of the concept. Phase II developed OSCAR software and demonstrated a fully working system. Phase III reconfigured OSCAR for use in F-16 air intercept training and conducted preliminary tests of OSCAR in conjunction with an Air Intercept Trainer (AIT) developed by the Armstrong Laboratory's Aircrew Training Research Division. A limited test was conducted with OSCAR using an AIT at the 58th Tactical Training Squadron, Luke Air Force Base, AZ. These data establish that pilots think that OSCAR was easy to use and would have a good training impact, probably for individual student practice and self-debrief. In the future, OSCAR may be applied to more complex tactical training used as a tool for training feedback research.

14. SUBJECT TERMS

expert systems
performance critique
performance feedback
performance measurement
training

15. NUMBER OF PAGES

58

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

CONTENTS

List of Tables

List of Figures

PREFACE

The objective of this effort was to develop an expert system and provide a prototype device to debrief users on performance of air intercepts. The Observing System for Critique, Advice and Review (OSCAR) is one of four related efforts to apply artificial intelligence methodologies to enhance aircrew training.

This research represents a portion of research and development at the Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division, in support of the Laboratory Research and Technology Plan, whose general objective is to increase combat readiness and job performance by developing and demonstrating more cost-effective ways of acquiring and maintaining new skills.

This research was conducted under Work Unit 1123-35-13, Observing System for Critique, Advice, and Review. Work was completed by VRC Corporation under Contract F33615-88-C-0005. Work Unit and Contract Monitor was Dr. Richard A. Thurman.

OBSERVING SYSTEM FOR CRITIQUE, ADVICE AND REVIEW (OSCAR)

INTRODUCTION AND SUMMARY

Summary

Expert critiquing uses expert system technology to produce a critique of a recorded engagement. The Observing System for Critique, Advice, and Review (OSCAR) is a test system which provides expert critique of performance as training feedback.

First, the technical approach requires an intelligent system which is capable of narrating an engagement in an expert fashion. Second, it requires an intelligent system which can critique the narration and, third, requires a capability to provide an effective debriefing to the users.

The purpose of this program is to test whether automated feedback based on expert critiquing can have utility during transition training.

A three-phase program was conducted: Phase I was devoted to initial development and a demonstration of the feasibility of the concept. Phase II was devoted to the development of OSCAR software and a demonstration of a fully working system. Phase III was devoted to a reconfiguration of OSCAR for use in F-16 air intercept training and the conduct of preliminary tests of OSCAR used with an Air Intercept Trainer (AIT) developed by the Armstrong Laboratory's Aircrew Training Research Division.

A limited test was conducted with OSCAR used with an AIT at the 58th Tactical Training Squadron, Luke Air Force Base, Arizona. The data were primarily subjective responses collected from 11 instructor pilots. These data establish that pilots think that OSCAR was easy to use, and would have a good training impact, probably for individual student practice and self-debrief. Additionally, some improvements and corrections to OSCAR were identified.

The test, however, was conducted with some of OSCAR's features restricted, with all processing conducted after an engagement was accomplished on the AIT (not in real time), and with a rule-base which was a strict translation of a tactical manual. Consequently, while the test results are favorable, little can be said about optimization of OSCAR's capabilities, or to applications which require complex performance narration and critiquing capability. In the future, OSCAR may be applied to more complex tactical training, or OSCAR may be used as a tool for training feedback research. Additional testing is recommended.

Technical Issues

The critiquing approach. The usual approach to computer decision-making is to design a system which simulates the expert's decision-making process. Such a system gathers data as an expert would and attempts to come to expert conclusions. If this approach is applied to air intercept training, the traditional approach attempts to tell the pilot (and instructor pilot) what to do. In contrast, a critiquing system assumes that the pilot has adopted a plan of attack or has flown an engagement. Rather than attempting to duplicate the decision process, the computer system critiques it, discussing the pros and cons of the approach compared to alternatives which might be reasonable or preferred. In this way, instead of arbitrarily advocating one approach to training, the computer system will let the pilot be its guide and tailor advice to the indicated thinking and plan.

A similar system is ATTENDING (Miller, 1984) which critiques a physician's plan for the use of anesthetics. Based on ATTENDING, the following benefits may hold for OSCAR: (a) The approach casts the computer in the role of the user's ally, rather than a potential competitor, (b) the user must think through the problem, (c) users can develop their own idiosyncratic style, (d) nuances, which can be difficult to quantify, can be treated, and (e) it leaves the major responsibility with the student and instructor, while the computer plays a secondary role.

Performance measurement issues. An automated measurement system must know what tasks are being performed. This, of course, is the essence of the performance narration capability required by OSCAR. While this is not easy for fixed profiles (e.g., instrument navigation), it is a major challenge for emergent combat missions such as air intercept.

A simple paradigm for measurement is to identify where tasks start and stop, and then measure the task in between these points. OSCAR must know what is happening, and must perform quantitative performance measurement even though the ultimate output may be subjective text. This is accomplished in OSCAR using a module called Smart Window-based Assessment and Measurement Interface (SWAMI).

To be useful, performance information must be provided at the time a task is completed, or very soon thereafter. The design of real-time or near-real-time measurement presents challenging issues. For example, the start/stop identification issues are alleviated if second-pass calculations on a full recording are possible. However, this is extremely difficult if the same start/stop identification must be done concurrent with live performance. SWAMI uses a back-looking, windowing approach with lagged logic to attack these problems.

Performance feedback. The effectiveness of the feedback information presented to the instructor pilot and pilot may depend on *how* the information is presented as well as *what* information is presented.

OSCAR provides feedback in the context of a graphical replay of an engagement, with critiquing occurring during the replay. Analysis of the critique information and associated causal factors is facilitated by allowing the user to examine the engagement at the time of the critique message and then explore preceding events. Key events from the performance narration can be used to rapidly skip to portions of the engagement of interest.

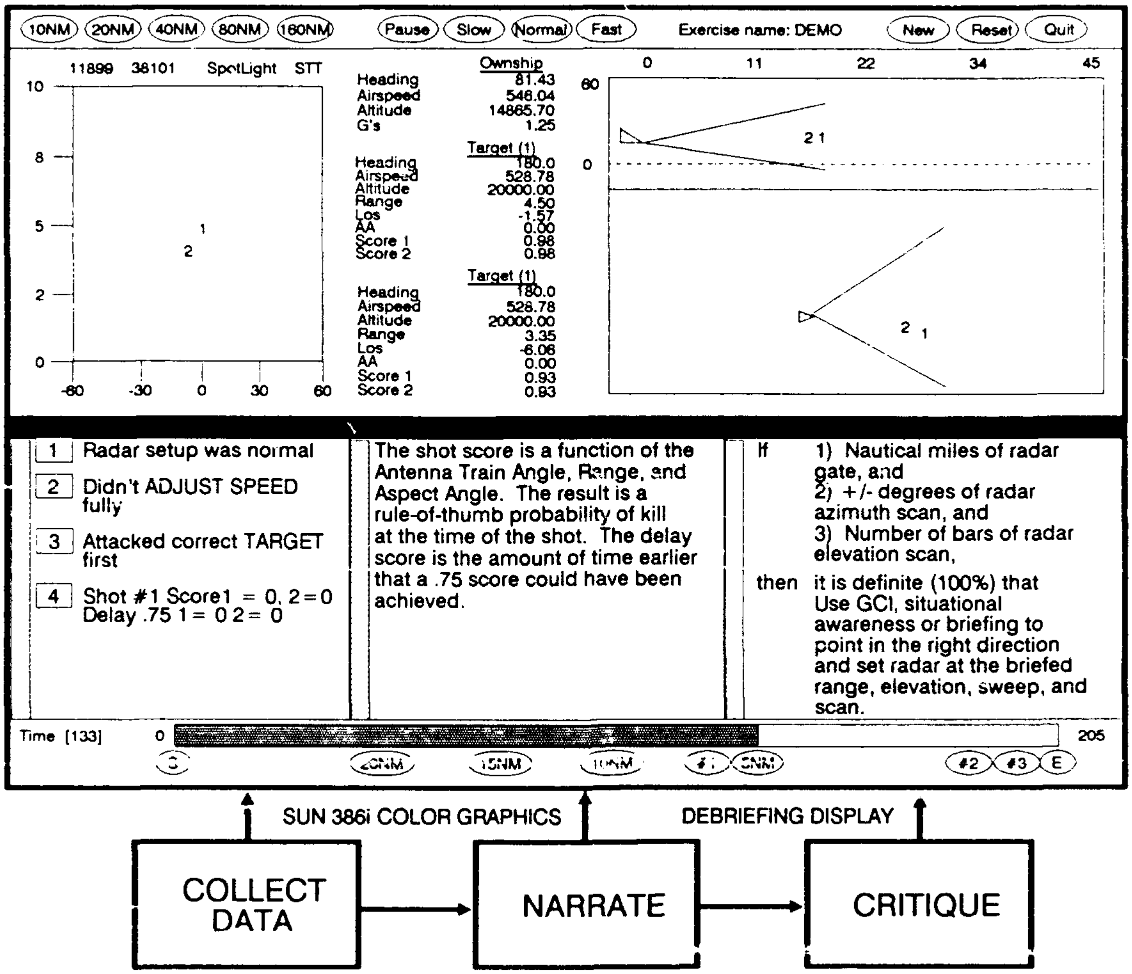
Description of the OSCAR/AIT Configuration

OSCAR has been configured for air intercept training and for use with the F-16 Air Intercept Trainer (AIT). The AIT is a joint venture between the Air Force Reserve, Tactical Air Forces, Air National Guard, and Air Force Systems Command. It is a result of an effort by Armstrong Laboratory to speed transition of improved training methods to users in a form which is affordable at the squadron level.

OSCAR consists of four main subsystems: (a) a data acquisition subsystem, (b) a narration subsystem, which computes and logs performance events which would be of interest to an expert observer, (c) a critique subsystem, which uses an expert system and a rule-based knowledge system to critique performance, and (d) a feedback subsystem which allows selective replay of performance along with display of the critiques and also allows the user to explore causal factors.

OSCAR

OBSERVING SYSTEM FOR CRITIQUE, ADVICE AND REVIEW FOR AIR INTERCEPT TRAINING



HOW OSCAR WORKS

TIME-SAMPLED AND DISCRETE DATA ARE COLLECTED VIA FLOPPY DISK OR DIRECT CONNECTION	A SMART WINDOW-BASED ASSESSMENT AND MEASUREMENT INTERFACE (SWAMI) CREATES A FILE OF MANEUVERS EVENTS AND MEASUREMENTS WHICH WOULD BE NOTED BY AN EXPERT OBSERVER	A RULE-BASED EXPERT SYSTEM MAKES INFERENCES ABOUT PERFORMANCE, RISK ASSESSMENT, PROBABLE CAUSES, AND ALTERNATIVE COURSES OF ACTION

PLAYBACK FEATURES

- o CONTINUOUS PLAYBACK WITH START/STOP SET BY TIME OR NARRATION EVENT
- o WHAT, HOW, AND WHY EVENTS HAPPENED, PLUS ALTERNATE COURSES OF ACTION
- o MULTIPLE GRAPHIC VIEWS: B-SCOPE, TOP/SIDE VIEWS WITH VISIBLE RADAR BEAM
- o AUTOMATIC CRITIQUE AS PLAYBACK PROCEEDS, PLUS HOW CRITIQUE FORMED
- o STORED DEMONSTRATION PLAYBACK FILES

Figure 1. Overview of OSCAR for air intercept training.

In this program, OSCAR was developed for use in air intercept training, as shown in Figure 1; however, the subsystems of OSCAR can be modified for use with other human performance applications.

As shown in Figure 1, the top third of the display is used to present a radar display and aircraft variables as they occurred during an exercise, and, additionally, a two-view display of three-dimensional space giving a God's-eye view. Consequently, the user can relate familiar views with the initially unfamiliar radar displays. Also, note that the radar beam is shown as a visible searchlight to make control of the radar beam apparent.

The bottom of the screen presents a horizontal thermometer-like display to show replay progress on a time-line. The replay can be restarted at any point in the exercise by pointing (using a mouse) at the timeline, or, since the exercise has been narrated, by pointing at event-codes at the sides of the timeline.

The middle of the display presents critique information and, if the user wishes, additional information on the basis for the critique and specific terms used in the critique. The user may make a menu selection to go to the point in the engagement where a selected critique triggers, and then may browse through the exercise to note related events and causal factors. The knowledge base for the critiques is based on the current air intercept tactics manual and was reviewed by available subject-matter experts.

An index of effectiveness was computed for each weapon delivery and, if there was a delay in taking a shot, the amount of delay was also computed. However, this feature was disabled for two reasons: (a) the weapons range model tables used for this computation are classified information, and (b) approximate computations could conflict with the launch zone information displayed on the heads-up display (HUD).

SYSTEM OVERVIEW

Expert Critiquing Systems

The usual approach to computer decision-making is to design a system which simulates the expert's decision-making process. Such a system gathers data as an expert would and attempts to come to expert conclusions. If this approach is applied to air intercept training, the traditional approach attempts to tell the pilot (and instructor pilot) what to do. In contrast, a critiquing system assumes that the pilot has adopted a plan of attack or has flown an engagement. Rather than attempting to duplicate the decision process, the computer system critiques it, and may discuss the pros and cons of the approach compared to alternatives which might be reasonable or preferred. In this way, instead of arbitrarily advocating one approach to training, the computer system will let the pilot be its guide and tailor advice to the indicated thinking and plan.

The critiquing approach is explored in this program. A system called Observing System for Critique, Advice and Review (OSCAR) was developed. A similar system is ATTENDING (Miller, 1984) which critiques a physician's plan for the use of anesthetics. The following are advantages which were attributed to ATTENDING, which also may be valid for OSCAR:

- The approach casts the computer in the role of the user's ally, rather than a potential competitor. The pilot, or instructor pilot, must not feel threatened.
- The user must think through the problem. The approach forces the user to grapple with the problem, and think through any difficult issues. This keeps the user centrally involved in the decision-making process.
- The problem area is very subjective. There are frequently several ways to approach a problem, and it is seldom that one approach is "right," and the others are "wrong." Each person develops his own idiosyncratic style, and would have little use for a computer-advisor which would not allow him to perform in the accustomed fashion.
- There are often nuances in a situation which are hard to anticipate and quantify, but which lead a person to lean one way or another in choosing an approach. Evaluation of these factors can be a very subjective process.
- It makes sense to leave the major responsibility with the student and instructor, and let the computer play a secondary role. The computer can help the student and instructor evaluate and optimize the approach taken.

There is merit in the type of focussed feedback that critiquing provides. A general discussion of alternatives is not likely to have the same impact as focussed feedback structured around the user's thinking. Nor would a somewhat arbitrary advocating a "best" approach have the same instructional benefit (except in basic training introducing the student to a new area). Implicit in this approach is that the user must have sufficient competence to generate a plan and to evaluate the system's critique; this, however, seems to be consistent with training at the transition or continuation level.

Performance Measurement Issues

If an expert critiquing system is to function based on air combat performance, there must be some means to automatically determine events, relationships, and variable values that an expert would need to know in order to assess performance. Consequently, the development of OSCAR requires facing some of the central issues in performance measurement, including automated performance measurement, near-real-time measurement, performance diagnosis, and measurement of expert performance (cf., Vreuls and Obermayer, 1985). These topics will be discussed briefly in the following paragraphs.

Automated performance measurement. Automating performance measurement permits measurement based on a greater number of factors than possible through direct observation. Data can be processed more quickly and personnel requirements can be reduced. However, it is difficult through automation to assess complex human performance with anything approaching the perspicacity of expert judgment. To do so requires intelligent measurement systems like OSCAR.

An automated measurement system must know what tasks are being performed. This is not easy for fixed profiles (e.g., instrument navigation), and is a major challenge for emergent combat missions such as air intercept. A simple paradigm for measurement is to identify where tasks start and stop, and then measure the task in between these points. OSCAR must know what is happening, and must perform quantitative performance measurement even though the ultimate output may be subjective text. To do this, we have expanded upon the development of a Smart Window-based Assessment and Measurement Interface (SWAMI).

Near-real-time measurement. To be useful, performance information must be provided at the time a task is completed, or very soon thereafter. For training, performance information loses a major part of its instructional value if not available at the time of student debriefing.

However, the design of real-time or near-real-time measurement presents challenging issues. For example, the start/stop identification issue is alleviated if second-pass calculations on a full recording are possible, and extremely difficult if the same start/stop identification must be done during live performance. SWAMI uses a back-looking windowing approach with lagged logic to attack these problems (see Figure 2). In part, SWAMI uses a technique in which data are buffered (called windows in the SWAMI design), and when sufficient data are collected to support a decision, measurement action can take place retroactively (lagged in time).

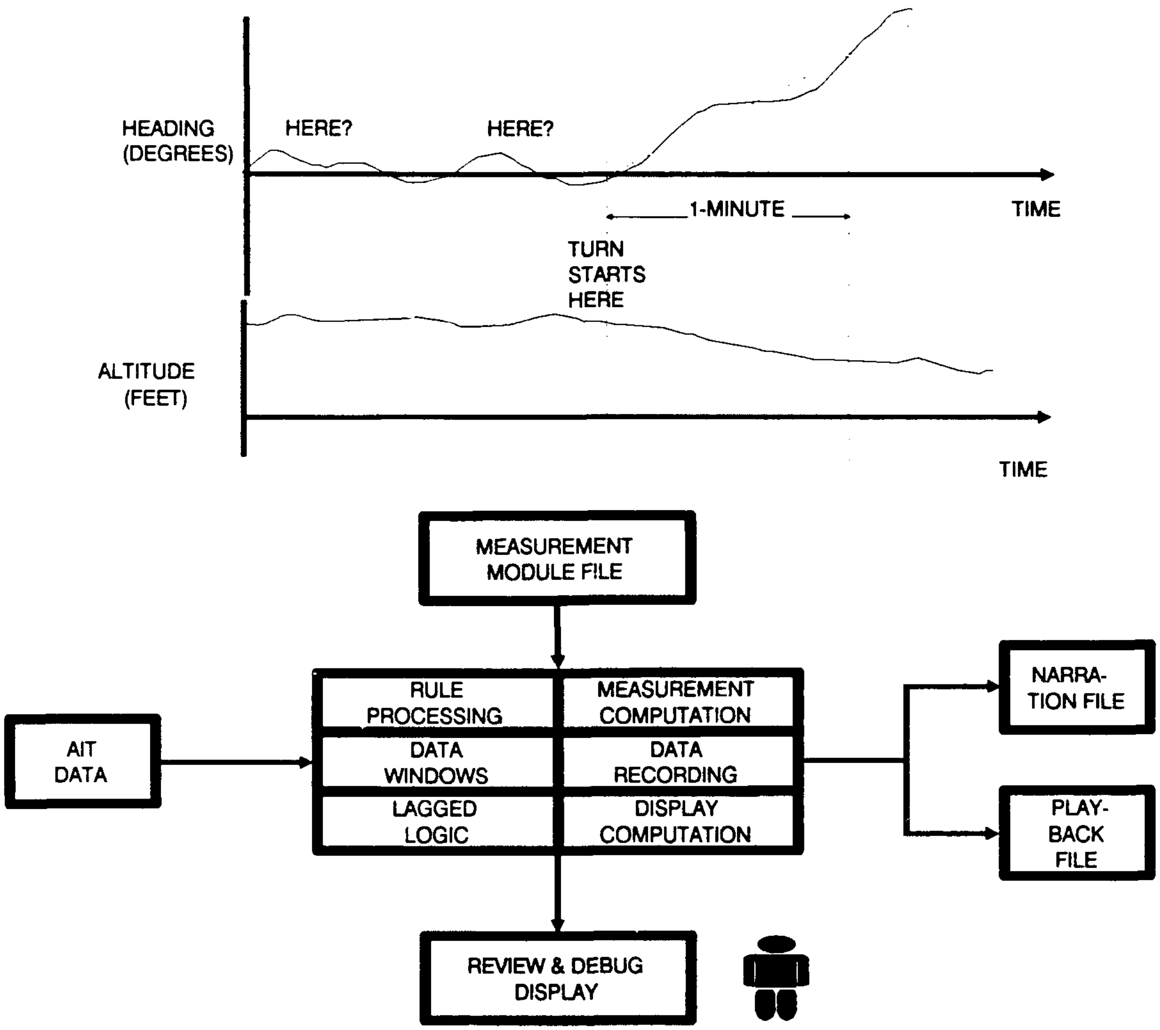
Performance diagnosis. The development of measurement for performance diagnosis is another major challenge which is addressed by OSCAR. The difficulty of diagnosis increases at a level below the obvious blunder where minor deviations from expected performance compound into an error. Here, measurement must be able to spot patterns over time, assess probabilities, and determine probable causes.

Measurement of expert performance. Closely related to the foregoing discussion is the issue of training people to become masters of their craft (e.g., well beyond the journeyman status). We tend to formally train to a minimum acceptable level of operational performance for practical reasons, leaving the further development of skill to on-the-job learning. Little is known about the differences between acceptable performance and the performance of experts.

PERFORMANCE NARRATION USING SWAMI

(SMART WINDOW-BASED ASSESSMENT AND MEASUREMENT INTERFACE)

PROBLEM: MEASURE ALTITUDE DEVIATION FOR ONE MINUTE AFTER INITIATING A TURN.
QUESTION: WHERE DOES THE TURN START?



- PROGRAMMABLE VIA MEASUREMENT MODULES
- RULES (IF...THEN...)
- DATA WINDOWS, LOOK-BACK, TWO-PASS PROCEDURES
- LAGGED LOGIC & COMPUTATIONS
- FLAGS, LOGIC BASED ON NONSIMULTANEOUS EVENTS
- REVIEW & DEBUG ENVIRONMENT

Figure 2. Overview of smart window-based assessment interface (SWAMI).

Knowledge-based Systems

Work on computerized expert systems grew out of earlier artificial intelligence work to make computers perceive, reason and understand. This work has progressed out of the laboratory and now there are many practical systems in operation. Knowledge-based systems can perform some of the most difficult decision-making jobs that include use of judgment, rules of thumb, and experience. This approach is of particular interest here because (a) there has been abundant success applying these techniques to automating human decision-making ability where the "spot-a-pattern, draw a conclusion" style of reasoning is used by experts, (b) it allows treatment of complex networks of knowledge which are overwhelming when other approaches are used, and (c) the knowledge base can be represented in a form which is understandable by the subject matter expert, which may lead to more sophisticated and unfiltered statements of expert knowledge.

A generalized form of a knowledge-based system to be used in this project is shown in Figure 3. We assume a system that is basically rule-based, which is often called a production rule system, rule-based system, or if-then system. The FACTS will be produced by automated performance

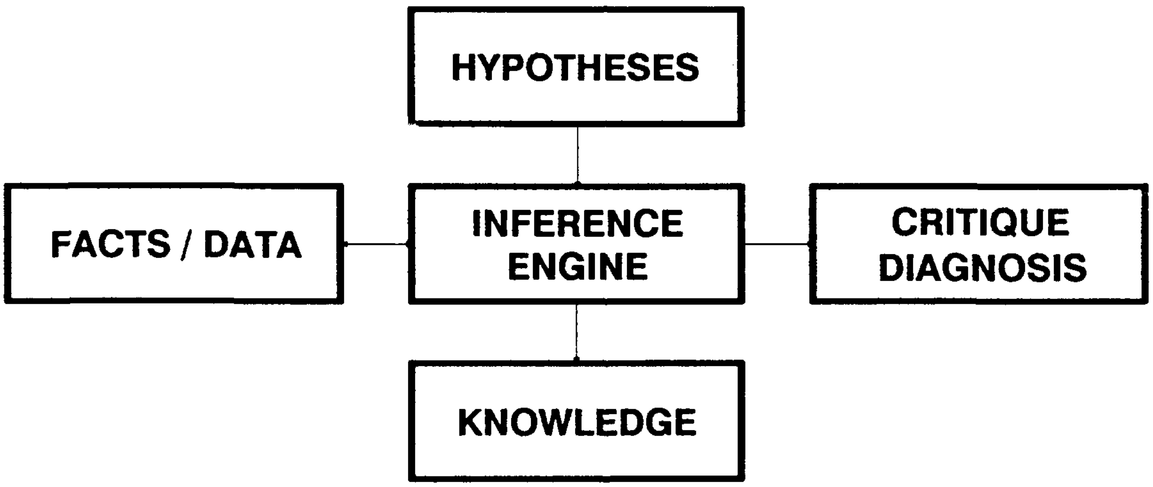


Figure 3. Block diagram of a knowledge-based system.

measurement based on time histories of engagements. The KNOWLEDGE or rules are statements of expert knowledge in the form: IF (premise 1)(premise 2)... THEN (conclusion 1)(conclusion 2)... The INFERENCE ENGINE is the control part which does the inferencing in either of two ways: (a) a forward-chaining method where the full set of rules is applied to the data, with conclusions remembered as new facts, and the rules are applied iteratively until no new facts are produced; and (b) a backward-chaining method where the rules appropriate to specific hypotheses are identified, and the facts which could lead to the desired conclusion are tested. It is generally accepted that the power of the knowledge-based system lies in the knowledge base and not, for example, in the inferencing procedure. It is therefore quite important that acquisition of expert knowledge be given a high priority.

To critique performance, one must first process engagement in near-real time and produce an expert narration of the engagement. When this is done, OSCAR will know, in detail, what has happened in the engagement.

Performance Feedback

The purpose of this program is to test whether automated feedback based on expert critiquing can have utility during transition training. The effectiveness of the feedback information presented to the instructor pilot and pilot may depend on *how* the information is presented as well as *what* information is presented (Scott and Fobes, 1982). As pointed out by Downs, Johnson and Fallesen (1987), "... create a situation where participants are active and interactive rather than passive as they are in the critique method. Research has consistently shown that active participation in a learning activity greatly increases the amount learned and retained, and that involvement can reduce one's resistance to recognizing one's own mistakes."

Research has investigated the form, abstraction and valence of performance feedback (Cusella, 1984; Downs, Johnson and Barge, 1984; Fisher, 1979; Ilgen, Fisher and Taylor, 1979; Ilgen, Fisher and Taylor, 1984; Jacobs, Jacobs, Feldman and Cavior, 1973; and Nadler, 1979). Investigation of form, that is, lecture or participation, indicates that participation (e.g., alternate question and comment sequences) decreases resistance, increases motivation, and allows detailed exploration. Investigation of abstraction indicates that detailed specific information should be provided during feedback. Investigation of valence (i.e., positive, negative, neutral, or mixed feedback) indicates that positive feedback is more credible to the recipients, whereas evaluative (opposed to praise) can be a major barrier to effective feedback.

Kaplan and Fallesen (1986) recommend the following categories of information for Army After-Action Reviews:

- What happened? Each key event should be described, including all characteristics and elements.
- How it happened? Key surrounding facts should be presented, including the preceding events and any related actions or events.
- Why it happened? The facts should be organized and inferences made with regard to probable causes.
- What alternative courses of action were possible? Other possible options should be presented along with tradeoff information.

As shown in Figure 1, OSCAR must include a display to present performance feedback, and this will require making design decisions about the depth and type of information in the critique data base, and the information displays included in the debrief display.

System Block Diagram

The system block diagram for OSCAR, as shown in Figure 4, consists of four major modules:

- Data acquisition--data is acquired which describes performance; this will include specific discrete events such as initial setup conditions, switch settings and mode changes, and digital values for performance variables (e.g., vehicle state variables, radar signals, geographic relationships) sampled at regular time intervals (e.g., N times per second).

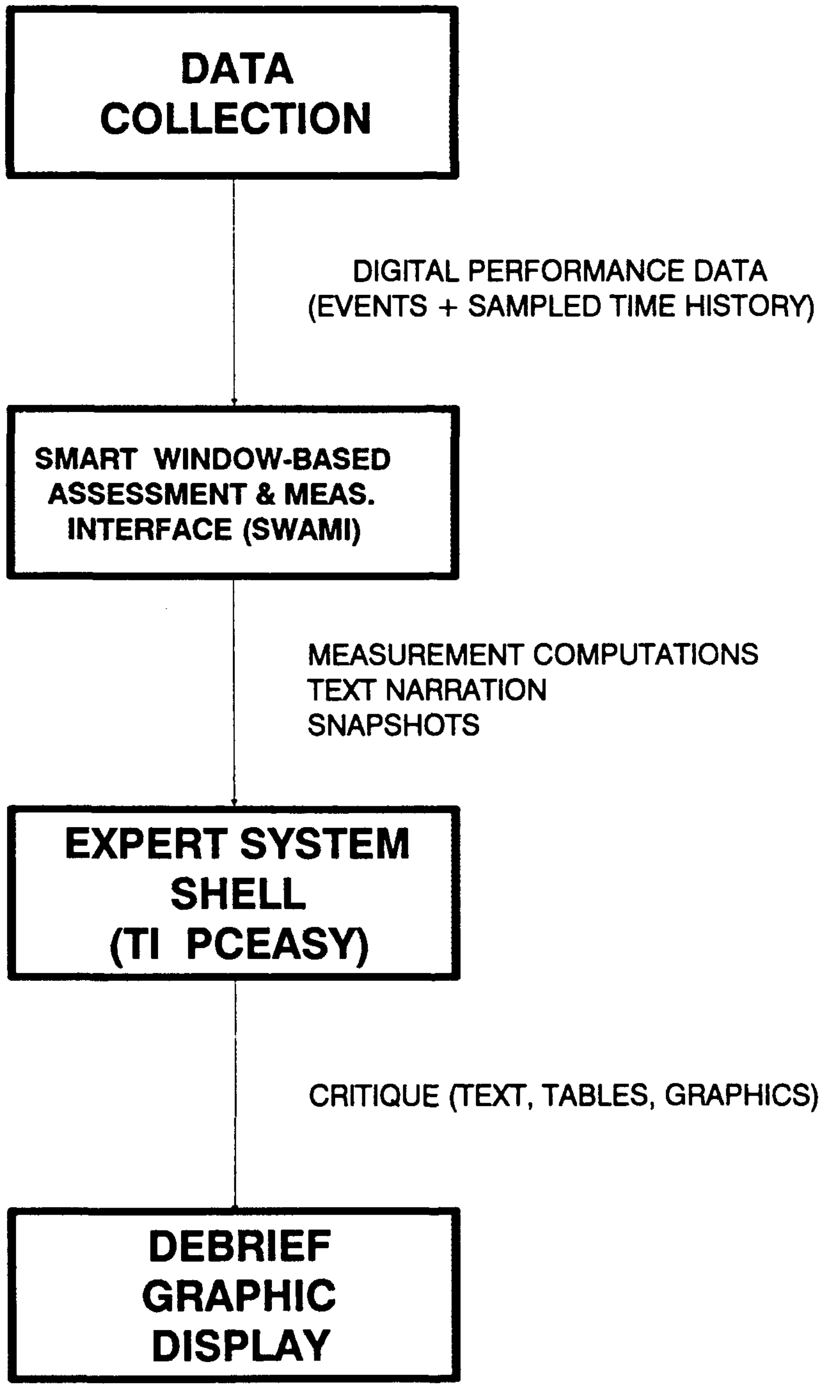


Figure 4. OSCAR system block diagram.

- Performance narration--SWAMI processes the digital data to produce a narration which is equivalent to that which would be noted by an expert observer (extended to the degree possible by the speed and accuracy of a computer). This may include results of computations, text noting th' specific tolerances have been exceeded, and "snapshots" of performance variables at specific times or conditions during an engagement (e.g., line of sight, aspect angle, range, velocity of closure, altitude difference, and speed difference at a specific point in an intercept).
- Performance critique, using a commercial expert system shell--the rules in the knowledge base of the expert system are applied to the performance narration to yield expert critiques for the given engagement. This may be in the form of text, tables or graphics.
- Debrief graphic display--presents feedback to an instructor pilot and pilot in context with detailed debriefing displays to allow the users to determine what happened, why it happened, determine possible causal factors, and establish alternative courses of action.

Each of these blocks is further discussed in further detail in subsequent paragraphs which describe OSCAR as it was configured for use in conjunction with the Air Intercept Trainer.

DESCRIPTION OF OSCAR

CONFIGURED FOR USE WITH THE AIR INTERCEPT TRAINER

First, during the program, a version of OSCAR was developed for a generic F-15 air intercept mission (Obermayer & Vreuls, 1989). This was done to develop and test key software modules and present a proof-of-concept demonstration. Subsequently, OSCAR was configured for testing for F-16 training using the AIT at the 58th Tactical Training Squadron, Luke Air Force Base, Arizona. The following paragraphs provide a description of this configuration (Additional detail may be found in VRC, 1990a).

It was determined that OSCAR was most applicable for training tactical considerations, and, the initial test configuration did not include critiques for basic intercept geometry. Also, it was determined that the advanced students at the 58th TTS would be ready for engagements against two nonmaneuvering enemies presenting a low-level, low-technology threat, but not for more complex tactical situations. Thus, only 1v2 setups on the AIT were used and OSCAR only provided critiques for these setups.

Figures 1 and 4 show that OSCAR consists of four major modules: (a) data acquisition, (b) performance narration, (c) expert critiquing, and (d) performance feedback (debrief). The performance feedback module also serves to integrate the four modules into a system with a common user interface. Each of these modules will be discussed in the following sections.

Data Acquisition

Data acquisition used existing capabilities of the AIT. The AIT includes a performance measurement subsystem programmed on a Zenith microcomputer. As a normal function, the AIT collected data on a pseudo-disk (use of high-speed, random-access memory to simulate a floppy disk drive). These data were copied to a floppy disk at the end of each exercise. The data included time, radar modes and variables, missile modes and variables, ownship position and flight variables, and position and flight variables for each target.

The data collection procedure dictated that all OSCAR processing took place after data transfer, and none could be accomplished during the exercise. Also, not all data available within the AIT were accessible; for example, heads-up display data and radar symbology were not available to OSCAR. Also note that all data are recorded as continuous variables sampled at a rate of 1 sample per second. Discrete switch settings, such as those on the radar control panel, are presented as continuous variables that change when the switches are moved to a new position.

Performance Narration

Performance narration is produced by a module of OSCAR called the Smart Window-based Assessment and Measurement Interface (SWAMI). The purpose of SWAMI is to process the data provided by the AIT and create a narration of the performance that is comparable to an expert observer.

Instructor pilots were asked to identify information that would be necessary for expert critiquing; that is, if they were to critique performance based on a printed set of measures and events, what would they need? They indicated that the following set of files should contain the necessary information:

- Discrete Events: switch settings, and the range and aspect angle at which they occurred.
- Intercept Summaries: for ranges from 30nm to 1nm, time, range, line of sight angle (LOS), aspect angle (AA), velocity of closure (Vc), altitude difference, speed difference. A separate file is generated for the first and second target attacked.
- Shot Summaries: for each weapon fired, time, range, LOS, AA, g, bank angle, target.
- Special Processing: scenario and target information, and special computations (not used) such as measures of min./max. beam altitude, maximum lateral separation, and all-aspect maneuvering index (AAMI) that estimates the effectiveness of a shot.

SWAMI was programmed to produce these files. SWAMI is capable of using data windows, back-looking and lagged logic to produce a performance narration in near-real time (that is, slightly lagging the real-time performance variables). However, since no AIT data were available to SWAMI until the end of an engagement, all SWAMI processing was performed after a floppy-disk transfer. Since near-real-time processing was not possible, a two-pass SWAMI narration was used because this was simpler than a one-pass procedure. The first pass, for example, accomplished the discrete event processing and shot summary. Then, in the second pass, the first-pass information can be used while second-pass processing is taking place. For example, the intercept summaries can be developed knowing which target was attacked first.

SWAMI contains a blackboard (that is, a special portion of computer memory) that permits data stored by a SWAMI procedure to be accessed by another procedure. SWAMI features are controlled by a Measurement Module (MM) file. SWAMI subsequently produces the narration files for each engagement.

Expert Critique

Discussions were held with personnel at the 58th TTS about the level of critiquing that they would apply to their students. They concluded that OSCAR critiques should be based strictly on the existing manual on mission employment tactics (Dept. of Defense, 1989). Therefore, the intercept rules closely conformed to this manual. As a result OSCAR presents a clear and consistent critique of AIT engagements. Note, that OSCAR can include conflicting expertise and debatable nuances in performance. The end result was a clear-cut "school-house" solution.

OSCAR incorporates commercially available expert system development software by Texas Instruments, called PCEasy. The knowledge base is divided into the following categories:

- Intercept-rules: baseline or head-on intercept based on TACM 3-3.

  Example (Rule 19)
  1) Speed advantage of ownship over the target is greater than or equal to 110 (simulator units), and
  2) Ownship is above the target,
  Then SPEED ADJUSTED not properly
  Source: TACM V, Pg. 4-82, Para. 4c4

- Detect-rules: radar setup and usage based on TACM 3-3.

  Example (Rule 05)
  If RCP setting: radar azi scan width is 10,
  Then RADAR AZIMUTH SCAN is not set OK.
  Source: TACM V, Pg. 4-76, Para. 1c3

- Tgt-select-rules: target selection table.

  Example (Rule 06)
  If 1) GCI is not available, and
  2) Formation of the enemy is BEARING/ECHELON, and
  3) 1) First target attacked is #1, and
  2) #1 is the closest target,
  or,
  1) First target attacked is #2, and
  2) #2 is the closest target,
  or,
  1) Aspect Angle stays greater than 160 degrees
  Then Selected correct target O.K.
  Source: TACM V, Pg. 4-81 and sort plan for single ship, low-threat bandit

- Narration-rules: identification of conditions used in the intercept, detection, and selection rules (e.g., "short lock," "tgt on the nose," "early-AA-less-than-120"). These rules use the SWAMI narration data to conclude that specific conditions exist.

  Example (Rule 36)
  If 1) Sign of altitude difference at 10NM is = 0, or
  2) 1) Sign of altitude difference at 10NM is , and
  2) Altitude difference at 10NM is = 1000
  Then Ownship is above target
  Comment: Rule allows for a "don't care" region of 1000 ft.

- Weapon-delivery rules: rules applicable to shot selection and timing; however, ultimately, no tactical rules were included in the initial application.

- Positive- and Negative-FB-rules: the time, conditions, and feedback messages presented to the student.

  Example (Rule 50)
  If 1) SPEED ADJUSTED is not proper
  Then 1) Display Negative Feedback Message #5
  Comment: Neg. FB Msg #5 is "Didn't adjust speed fully"

- System-rules: translation of scenario initial conditions and execution of operation system features.

  Example (Rule 60)
  If (unconditionally true statement)
  Then 1) DONE
  2) call a DOS program named QUIT.COM
  Comment: DONE is the last goal attempted

Debrief System

OSCAR uses computer graphics to provide a feedback display on a SUN 386i color graphics workstation, as shown in Figure 4. The intent is to present OSCAR critique information in the context of a full playback of performance, and to allow the user to explore cause-and-effect relationships pointed up by the critique messages. The following features are included:

- B-scope display. The information presented on the B-scope is available for integration with x-y-z display information and the critique information.
- X-Y-Z display. X vs Y and X vs Z information is presented in two views. The coverage of the radar beam is made visible and this may clarify critiques to new students. This display is meant to resemble a performance feedback display that is available in some versions of the AIT.
- Aircraft variables. Aircraft state variables and intercept geometry variables are displayed in full precision for integration with the critique information and other performance displays.
- Critique windows. Three windows are available to present critique information and related information. At the proper time, critique messages will appear in the window, and then scroll up as new critique messages are triggered. By pointing to a critique, using the mouse control device, a menu with three options is displayed to the user. One option is to reset the playback to the time at which the critique message was displayed. A second option is to have an explanation of the critique displayed in the second window. The explanations contained text taken from the tactical manuals used to develop the critique rules. The critique messages were designed to be very short, allowing many messages to be displayed at one time. A third option is to present the rule upon which the critique was based; this information is displayed in the third window. As a result, the user may see at one time (at the proper time in the playback), the critique, explanation, and the rule used in OSCAR's rule base.
- Playback control. Buttons and a time line are provided for user control of the playback. The user can use the buttons (selected with a mouse) to control the playback in the manner of a conventional tape recorder; or, the user can select a specific point in the mission from the time line. The time line shows both time and specific narration events.

These displays, of course, require that the user is sufficiently knowledgeable in the task to be able to interpret the information presented. A detailed presentation of *what* happened during the engagement is available to the user. Additionally, OSCAR provides the user with control

to find key events and to examine the preceding portion of the engagement; thus, the user has information on *how it happened*. *Why it happened* and *alternative courses of action may be inferred by the student. However, explicit information on these topics may be required from either an instructor pilot, or from additional rules in the critique knowledge base.*

Figure 5 shows the log-on screen that the user uses to begin an OSCAR session. The user may load new engagements and select engagements for debriefing by selecting from a menu of engagements keyed to the user's initials. The log-on screen integrates the four major modules of OSCAR. That is, selection of the load engagement function automatically sequences the data load, narration, and critique functions.

The procedures for using OSCAR, in the form of a guided tour of features, are presented in Appendix A. The user is prompted to use OSCAR features for (a) engagement selection, (b) playback control, (c) control using the narrated timeline, and (d) use of the explanation/rules/go-to time menu. The material in Appendix A is intended to be used by any first time user of OSCAR.

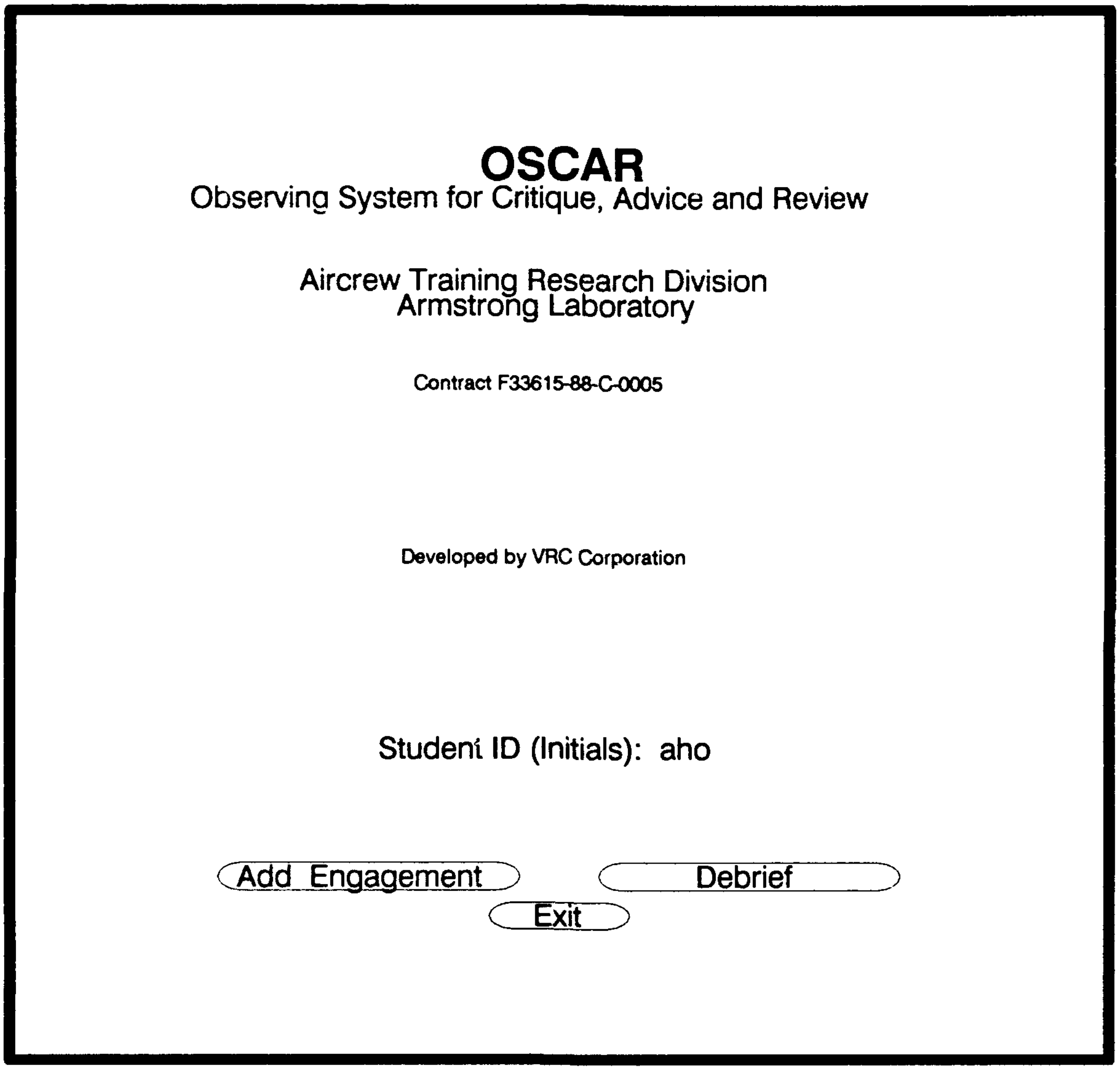


Figure 5. OSCAR log-on screen.

PRELIMINARY TRAINING TEST

This section presents the results of a preliminary test of OSCAR using students and instructors made available through the 58th Tactical Training Squadron, Luke Air Force Base, Arizona. OSCAR and testing methods are described and the results presented. The findings are discussed in terms of potential design improvements and an assessment of the utility of OSCAR for air intercept training.

OSCAR was specially configured for this test in correspondence with guidance from representatives of the 58th Tactical Training Squadron. The previous section describes the configuration. This configuration assumes that advanced students will be used, who have mastered basic air intercepts, and who are ready to try intercepts against two low-threat intercepts. The critique knowledge base was a translation of the existing tactical manual into a set of rules. Since the available data did not include heads-up display dynamic launch zone information, and weapons range model information is classified, no weapon's delivery critique was included.

Subjects

All subjects were scheduled through the 58th Tactical Training Squadron, Luke AFB, Arizona. Table 1 summarizes the subjects' F-16 flight time and years in service.

TABLE 1. Summary of Subjects' Experience

TYPE SUBJECT	NO.	F-16 FLIGHT TIME		YEARS IN SERVICE	
		MEAN	RANGE	MEAN	RANGE
LOW-TIME STUDENTS	3	3.03	3 - 3.1	3.83	1.5 - 5.0
HIGH-TIME STUDENTS	3	95	85 - 450	7.8	6.0 - 10.0
INSTRUCTOR PILOTS	11	704	170 - 1000	12.1	7.0 - 18.0

Table 1 shows that three types of subjects were used. The low-time students have little radar or intercept experience. The high-time students have a good foundation in radar and air-air intercepts, but are advancing beyond their course work in attempting a multiship tactical intercept. Most of the subjects were instructor pilots with extensive experience.

Experimental Procedure

A package of assessment materials was given to each subject (see Appendix B). The package contained written instructions, and there was minimal oral interaction with the subjects. The assessment package included a self-guided tour of OSCAR features that presented each feature and prompted the subject to try the feature on OSCAR.

After finishing the self-guided tour materials, each subject completed an experience questionnaire and rated OSCAR on four scales of bipolar adjectives. Subsequently, each subject performed at least two intercepts on the Air Intercept Trainer and then self-debriefed these exercises using OSCAR. Following that, the subjects again rated OSCAR on the same four scales, and then completed a questionnaire.

Although each intercept included two targets, the subjects were not required to engage the second target. Normally, after firing on the first target, the second target should be within visual range; however, the resolution of the AIT display did not always allow visual acquisition. So, the subjects were instructed to try to find the second target, but not to persevere if this proved to be impossible.

Experimental Measurement

Three types of measurements were administered: (1) Bipolar, semantic differential rating scales, (2) questionnaire, and (3) frequency of use of OSCAR features.

Bipolar (Semantic Differential) rating scales. Bipolar adjective scales represent a well-established method to measure reactions to a stimulus (Osgood, Suci, & Tannenbaum, 1957). Ratings on bipolar adjective scales tend to be correlated, and Factor Analysis concludes there can be three major stable dimensions, which have been labeled Evaluation (E), Potency (P) and Activity (A). Additional research on communication has elicited a fourth factor, labeled Style (S) (Carroll, 1960). It is possible that some or all of these bipolar scales will correlate with the E scale. But the initial design allowed for the possibility that the four scales are independent. In designing an instrument based on the bipolar adjective approach, it is desirable to select adjectives for which there is empirical evidence that they (a) can be interpreted as adjectives with opposite meaning, and that (b) can correlate exclusively with one of the E, P, A, or S factors. The folllowing bipolar adjectives were selected based on these criteria:

- Evaluative: valuable-worthless
- Potency: deep-shallow
- Activity: fast-slow
- Style: precise-vague

These scales were used twice: first, just after the subject had received instruction on the use of OSCAR, and, second, just after the subject had used OSCAR to debrief two intercepts. Thus, one can examine the shift in ratings due to experience with OSCAR besides the individual ratings.

Post-study questionnaire. The post-study questionnaire did not attempt to elicit answers to specific questions which could not be clearly identified in an exploratory study. Instead, the questionnaire caused the subject to conduct a broad review of OSCAR features across a variety of experiences.

The questionnaire consisted of three parts:

- a review of OSCAR by (a) information content, (b) presentation of information, and (c) capability to explore and analyze each phase of the intercept (sort/target, intercept, and weapon delivery). Note, because the All-Aspect Maneuvering Index (AAMI) score included in OSCAR would be classified if it were accurate and agreed with the heads-up display, the AAMI score was disabled. The subjects were asked for a good-bad rating on a five-point scale and any pertinent comments.

- a review of each OSCAR display and control feature: (a) log-on screen, (b) playback control buttons, (c) intercept playback displays, (d) critique messages, explanations, and rules, and (e) narrated time-line control. The subjects were asked for a good-bad rating on a five-point scale and any pertinent comments.
- a series of open-ended questions about (a) procedures for using OSCAR, (b) training effects, and (c) best and worst features of OSCAR.

Frequency of use of OSCAR features. The implementation of OSCAR included a feature for recording each control action by the user, and the time of activation. In the analysis, each control action was counted.

Results

The following paragraphs present a summary of the results (for additional detail, see VRC, 1990b). The data for one instructor pilot was removed from the analysis because it was clear that, due to some miscommunication, his ratings and comments were in the context of the AIT and provided no information on OSCAR. Because the sample size is small, and the study is preliminary and informal, no statistical analysis was conducted.

Bipolar adjective scales. Mean ratings on the four bipolar adjective scales, for ratings just after the introductory guided tour of OSCAR features, and for ratings after using OSCAR to debrief intercepts flown on the AIT, are presented in Figure 6. Data are presented separately for low-time students (N = 3), high-time students (N = 3) and instructors (N = 11).

All ratings showed a strong positive attitude toward OSCAR. Further, for the instructors, the positive attitude was stronger after experience with OSCAR.

Good-bad ratings on questionnaire. Good-bad ratings on a five-point scale were analyzed for (a) the questionnaire on Feedback by Engagement Phase, and (b) the questionnaire on Design Features. All ratings were high except for weapon delivery items. These ratings reflect that weapon delivery capabilities were disabled.

Frequency of use of OSCAR control features. Table 2 lists the average number of times per intercept that each control feature was used.

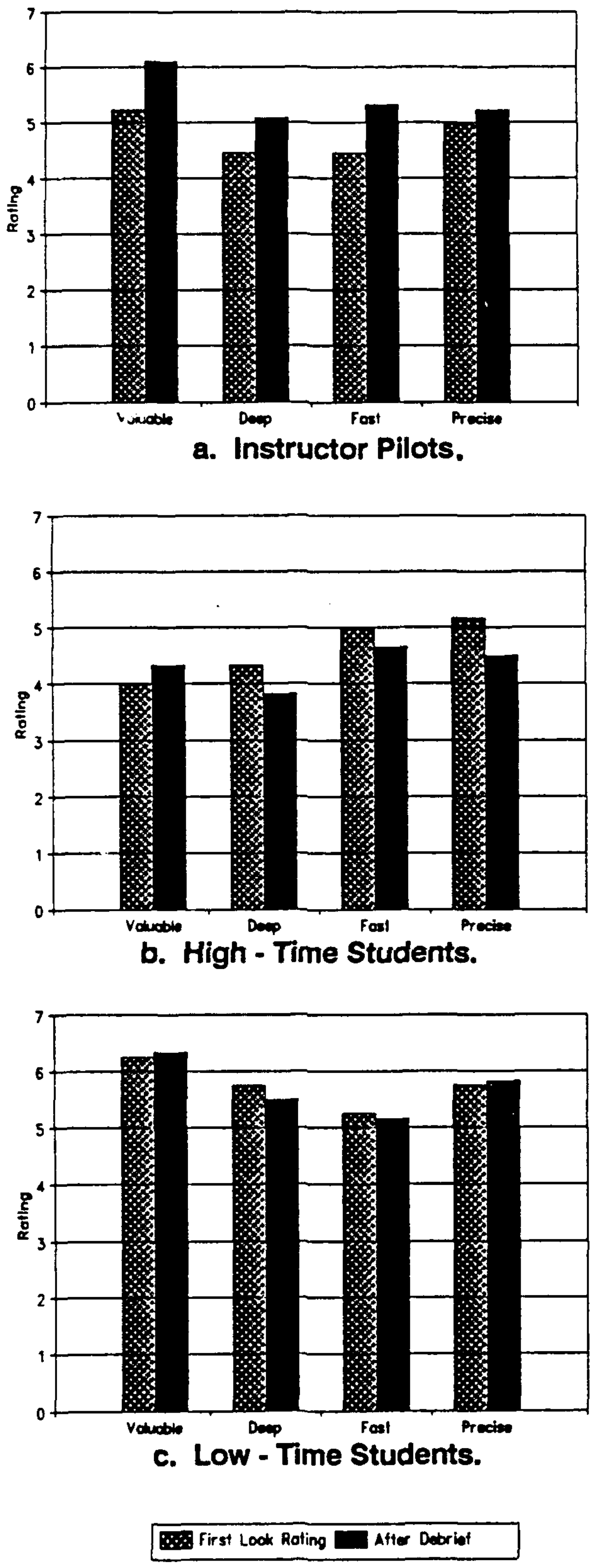


Figure 6. Mean ratings on the four bipolar adjective scales.

TABLE 2. Frequency of Use of OSCAR Control Features
(Average Number of Uses per Intercept Debrief)

FEATURE	LOW-TIME STUDENTS	HIGH-TIME STUDENTS	INSTR. PILOTS	TOTAL
BSCOPE SCALE	0.40	0.33	0.56	0.49
PLAY: PAUSE	1.00	0.67	1.03	0.95
PLAY: SLOW	0.70	0.42	0.31	0.39
PLAY: NORMAL	1.00	1.50	2.13	1.82
PLAY: FAST	0.90	1.33	1.21	1.18
CONTROL: NEW	0.10	0.08	0.36	0.26
CONTROL: RESET	0.30	0.17	0.33	0.30
CONTROL: QUIT	0.70	0.75	0.64	0.67
MENU: EXPLAN.	2.30	0.42	1.79	1.61
MENU: RULES	1.20	0.83	1.08	1.05
MENU: GO-TO-TIME	0.40	0.42	0.44	0.43
POINT TIMELINE	1.20	0.50	0.92	0.89
NARR. BUTTONS	0.80	0.42	1.18	0.97
TOTAL EXERCISES	10	12	39	61

These data are for only a few intercepts for each subject; thus, there was insufficient time for the subjects to develop either a clear preference for the use of specific control alternatives, or a specific strategy for using OSCAR features.

Comments. The following paragraphs summarize the written comments made by the subjects:

- An error discovered. Some early data pointed out that the critiques indicated the proper adversary had been targeted, when the opposite was true. A last-minute revision to the sort plan rules eliminated wrong targeting as an issue in a head-on attack. For the intercepts in question, a portion of the intercept resembled a head-on attack. The rules were revised to (a) make a stricter and more thorough test of the conditions for a head-on attack, and (b) to remove positive feedback about correct target selection from head-on engagements. These changes were made at the end of the first day, and no other problems with target selection were noted.
- Incorrect radar beam display. The coverage of the radar beam, as displayed on OSCAR, was based on early-model information; so, some subjects noted a difference between OSCAR and the AIT radar.
- Weapon delivery critique. Weapon delivery data and critique should be added to OSCAR. It was suggested that a "what-if" capability be added to allow exploring options in weapon delivery. A simulated "fly-out" of the missile also was suggested.
- Nuances, variations in technique. The rule base used in OSCAR is a strict translation of the current tactics manual, and provides a single "schoolhouse" critique of performance. However, the design of OSCAR includes the potential for alternatives, opposing points of view, and tradeoff analysis. Some suggested that OSCAR was too narrowly focussed on one way to accomplish the job.

- Display size. OSCAR provides fixed windows on the terminal screen for each sub-display. Some recommended a "zoom" capability to increase display size and scale (especially within 5 nautical miles range).
- Removal of information. Explanations were left on the screen after being requested. Consequently, the information might be inappropriate at a later time when some other critique is displayed.
- Additional information available on cockpit displays. OSCAR displays were limited by the information transmitted to it from the AIT. Some suggested that additional information should be acquired and displayed on OSCAR. In particular, radar symbology such as "bugged target" and target aspect, and a heads-up display should be added.
- OSCAR connected to real-time intercept. Subjects suggested that OSCAR should allow doing the intercept again, starting over at the point of an error.
- Explanations. The information for further explanations was taken directly from the manual (Dept. of Defense, 1989); however, the clarity of some explanations was questioned.
- OSCAR was easy to learn and use. Apparently users found OSCAR easy to use after the brief guided tour (Appendix A).
- Display of rules. Some subjects questioned the interpretation of some displayed rules (used to trigger critiques).
- Additional output. Addition of a hardcopy printout (of critique messages) would reduce the amount of interaction with the OSCAR computer.
- Procedures for using OSCAR. A wide variety of procedures for using OSCAR was tried. However, the users debriefed only two full engagements and did not have sufficient experience to come to firm conclusions.
- Training effects. A universal conclusion was that OSCAR would have a good training effect. However, there were a variety of opinions about how training would be impacted. It was generally thought that OSCAR would be best for the "beginning," "basic," "initial," "introductory" or "new" student.
- Individual vs. instructor-student debrief. Many subjects seemed to favor use of OSCAR for self-debriefed practice; however, both student-alone and instructor-student debrief use, as well as other points of view, were expressed.
- OSCAR's best feature. A variety of "best" features was identified, including: flexibility, potential for stand-alone training, debrief displays, integrated analysis, instant feedback, ease of use, and rule-based critiques.
- OSCAR's worst feature. A variety of "worst" features was identified, including: specific critiques vs. individual technique, a bit slow, need radar and HUD information, visual display, displays too small, no shot analysis, rule explanations, format of some information.
- Errors and bugs. Some errors that were mentioned: targeting incorrect (already fixed), need missile Probability of Kill information and shot analysis, need a more extensive rule base.

DISCUSSION

The following paragraphs present issues that arose either during the development of OSCAR or during preliminary training tests. The development of OSCAR required design decisions for which ready answers did not exist. The training tests also pointed up additional issues besides the limitations of these tests in addressing the full potential of OSCAR.

Critiquing *What* and *How*

Performance feedback should include information on (a) what happened, (b) how it happened, (c) why it happened, and (d) alternative courses of action. Information on *what* happened and *how* it happened is included in OSCAR debrief displays and critique messages; these categories of information are discussed first.

Message format. The design of OSCAR critique messages attempted to maximize the number of messages that could appear on the screen at one time. If all the critique messages could appear at one time, the user could make a general assessment and pick the messages that should be investigated first. Otherwise, if all could not fit on the screen at once, the user would have to take action to scroll through the messages. One cannot be sure that such short messages will provide sufficient feedback, even though instructor pilots checked the content of each message, so further explanation was presented in another window. The text for the further explanation associated with a critique message was taken from the tactical manual (which was used for development of critique rules). Thus, all relevant parts of the tactical manual were available for on-line interrogation. Additionally, the user also could access the rule associated with a critique message, and have this presented in a third window.

The format used for critique message presentation seemed to be acceptable to the test users. Only a few comments about message clarity were received, but these related to text taken from the tactical manual.

While subjects used the option to display rules, it is not clear how this may have contributed to critique feedback. Possibly, rules were displayed to check the validity of the critique message, and therefore may have contributed to user acceptance of the feedback. Only the top-level rule that triggered a critique was displayed; to be maximally useful, the display should allow tracing through the hierarchical layers of rules that may be involved in triggering a critique.

The use of short messages did seem to insure that all short critique messages did get displayed simultaneously. However, the addition of weapon delivery critique probably would add enough messages to exceed the display capability, and require that the user scroll through the messages to be sure that all were read.

If multiple points of view are to be presented, separate windows for each point of view may be required to avoid confusion. Explanations and rules may then require display on windows that temporarily overlay portions of the screen.

Selecting time for presentation of critique. Besides determining the critique messages to be displayed to the user, it was also necessary to determine *when* in the playback to present the message.

For out-of-tolerance conditions, it was clear that the corresponding message should be presented when the relevant variable was out of tolerance. Ordinarily, the message was associated with tests conducted at various distances to the target. The time for message display was selected as the farthest distance for which the variable was out of tolerance.

However, for some critique messages, the time to display the message is less obvious. For example, it may not be clear that the wrong target was selected until a weapon is fired at the wrong target. The critique message should be displayed at an earlier time when the mental decision was made.

While when-to-display decisions are embedded in the current knowledge base, it may be desirable in the future to create a separate portion of the knowledge base that includes the appropriate instructional knowledge.

Comprehensive, high-fidelity display. OSCAR displays were simplified versions of the actual cockpit displays. It was thought necessary, and perhaps even desirable, to reduce some of the complexity of the cockpit displays. It was believed that the simplified displays would allow the user to re-create the engagement in sufficient detail. For example, OSCAR did not include information from the HUD, nor did the radar B-scope include all symbols (such as the "bugged targets" that provide information on target priority).

It may be argued that the OSCAR displays were adequate for most of the critique and diagnosis involved in the low-level 1v2 tactics used in the preliminary tests. Nevertheless, it is clear that most of the instructor pilot subjects wanted a closer approximation of the cockpit displays. At one level, there was difficulty in dealing with displays that were different from their customary displays. On the other hand, detailed diagnosis of performance difficulties may require display of cockpit information that was not obvious to OSCAR display designers. As a rule of thumb, it may be better to err in the direction of overdesign for playback displays.

Critiquing *Why* and *Alternative Courses of Action*

OSCAR provided explicit information on *what* happened during a performance anomaly, and directly related information on *how it happened*. However, performance feedback also should include information on *why it happened* and *alternative courses of action*. At the level of many critique items in the OSCAR preliminary test, e.g., insufficient altitude difference, the information may be somewhat trivial; but, it is far from trivial for many questions about targeting and weapon delivery.

Such information could be added to the OSCAR knowledge base, resulting in the display of information on probable causes and tradeoffs associated with alternative courses of action. The primary challenge is to acquire sufficiently detailed knowledge from experts and to present the information meaningfully.

Creating opportunities for exploration and discovery. To the extent that there is not explicit knowledge for causes and alternatives in OSCAR, the user may be required to extract such information from the playback displays. Of course, this assumes that the user has sufficient training to permit diagnosis and prescription; that is, the user is a sophisticated student, or an instructor is present. Given this, the user should be encouraged to explore the playback information. OSCAR gave the user the capability to find the place where a critique message occurred, and then to move easily through the engagements to explore sequences of events and causal relationships.

Nonchronological playback. Although playback similar to a tape recorder was available (at slow and fast speeds), the nature of the digital database allowed the user rapid access to any point in the engagement. Then OSCAR displays could be reviewed at the selected point. Many of the instructor pilots used OSCAR by setting the playback to the end of the engagement or near to weapon delivery. Then they could note the critique messages, and begin exploring the engagement in the region of the more critical messages. In general, such exploration flowed backward, from the end or a critical point in the engagement, sampling forward to see events that led up to the selected point.

The narrated timeline. The narrated timeline was a key feature of OSCAR that facilitated exploration and discovery. By pointing to the timeline (using the mouse), the engagement could be immediately reset to that point in the playback. Because the engagement was automatically narrated during the critiquing process, the timeline could be labelled in terms that were meaningful to the user. As a result, without knowing the actual time, the users could select an event by pointing to a label next to the timeline (e.g., 5NM range, Shot #1). This feature seemed to have a distinct effect on the way users conducted OSCAR's "playback."

Critiquing Approaches

Valence and quantity of feedback. There is much emphasis in the literature on presentation of positive feedback. However, it seems necessary to point out deficiencies using negative feedback. If both positive and negative feedback is presented for performance that has many features to be critiqued, then there will be many messages presented. There will always be a message, either positive or negative, for every feature of performance. For complex performance, this quantity of messages could overwhelm the student and dilute the effectiveness of feedback.

The approach taken, in the version of OSCAR used for preliminary testing, was to provide negative feedback for any performance deficiency, but to provide positive feedback only for high-level integrated units of performance (e.g., the entire intercept, selection of the appropriate target).

Critique content, relation to instruction. Much of the content of the tactical manual, upon which the test version of OSCAR was based, was in terms of specific choices and tolerances for variables. In such a case, the critique content is very clear. Presumably, the critique should support the instruction given in the classroom and aircraft. The critique can reinforce and continue the instruction, and design of critique message content and supporting explanation should attempt to augment instructional goals. Furthermore, one should consider critique at multiple levels; for example, besides the specific tactics and errors in execution, the critique also can comment on total strategy.

Real-time vs. after-engagement feedback. Immediate feedback is desirable. Therefore, the effectiveness of delayed feedback will depend on the sufficiency of the students' recall (with the aid of performance playback, such as that used in OSCAR). Thus, there may be distinct advantages to using OSCAR in near-real time, giving feedback during the engagement. OSCAR contains the data processing capability to do this but, during the tests, could only obtain post-engagement data on a floppy disk.

There are at least two issues that should be addressed if OSCAR is to be used for near-real-time feedback. One issue is the potential for feedback interfering with the student's performance; or, feedback may be ignored because the student is preoccupied. It may be possible to use a technique in which some key feedback occurs during the engagement and other feedback is delayed for a post-engagement playback. A second issue is whether the logic associated with critique of specific performance can be executed in a timely way. For example, feedback about target selection should be given when decided, but the selected target cannot be identified until there is an over action.

Nuances and alternative points of view. The version of OSCAR that was tested was based on a specific "schoolhouse" solution for tactical air intercepts. This is the requested capability and this seemed appropriate for preliminary testing. During basic intercept instruction the students are taught a specific nominal approach and then they must demonstrate proper execution. Nevertheless, some instructor pilots commented during the preliminary tests that the rules should be modified to allow for individual techniques. However, advanced levels of training may require automated feedback to consider nuances in performance and, where knowledge is uncertain, consider conflicting points of view.

What-if and Alternate Courses of Action (ACOA). Theoretically, the use of an expert critiquing to provide information on alternative courses of action has been used in the medical community for administration of anesthetics (e.g., ATTENDING (Miller, 1984)). An addition to the OSCAR knowledge base can be made for this purpose, but this requires acquiring more knowledge from experts. Then, OSCAR can initiate a dialogue with the student, pointing out performance alternatives and the associated advantages and disadvantages. Furthermore, it should be possible to save the conditions at a specific point in an engagement. Then one can use these as initial conditions for the simulator, allowing the student to complete an engagement in a different manner.

Acceptance of computer feedback. OSCAR was very well received during the preliminary tests done at the 58th Tactical Training Squadron. As a result, it can be stated that automated feedback based on expert system technology can be acceptable for tactical training. On the other hand, it should be noted that the feedback was based on a strict application of the existing tactical manual. It is probable that acceptance was based on careful checking by the instructor pilots that the rules conformed with the manual. Further, it seems that impersonal computer-derived feedback may have advantages for self-administered practice, providing a privacy to make mistakes without retribution or embarrassment. Nevertheless, if OSCAR is to be applied to performance for which nuances and conflicting points of view are a major feature, the issue of acceptance should be carefully examined.

Training Issues

Student-only practice, student/instructor feedback. The preliminary tests with OSCAR used only a one-person debrief: either a single student, or a single instructor pilot who, in part, was simulating a student. The subjects judged that the addition of OSCAR to the AIT created an environment for good self-practice, that would allow the AIT to be used well in a way that would reduce the load on the instructors. With OSCAR, the student can fly the AIT until assured that performance meets school requirements.

OSCAR testing has not included use with both a student and instructor present. The presence of the instructor would undoubtedly enhance the feedback, and help the student in determining how to improve. However, the design of OSCAR feedback could be changed, because now the goal would be to augment and standardize the feedback provided by the instructor. For performance nuances and conflicting points of view, the instructor could provide an aid to the student in filtering, ranking and integrating the information presented.

Assessment and prescription. If an instructional knowledge base is to be added to OSCAR, knowledge would be added to help the student in assessing the performance and deciding what to do differently during the next engagement. That is, feedback also could include information on *what is expected* and *what might be improved*. Additional rules could be added to point out that some deficiencies are relatively minor and to be expected at the present stage of training. Others, however, are major and require immediate emphasis. Beyond this, OSCAR could be augmented with a cumulative performance data base, permitting direct comparison with other students at the same point in training.

CONCLUSIONS

The conclusions presented in the following pertain either to (a) the degree of achievement of the program objectives, or (b) directions for later work.

Program Objectives

The principal program objective, following demonstration of the feasibility of the concept, was to determine if automated feedback can have utility during transition training.

The question of feasibility may be treated in four parts corresponding to the four major modules of OSCAR: data acquisition, narration, expert critiquing, and feedback presentation. Each of these, with perhaps some revision as noted during testing, was shown to be capable for basic tactical intercept training. The narration module (SWAMI) has greater capability than required for tactical intercepts. Near-real time narration of continuous performance is possible, although the capability for spatial inferencing associated with within-visual-range maneuvering has not been tested. The commercial rule-based expert system appears to be extendable well beyond the current capability; however, knowledge acquisition may be a difficult and labor-intensive process. The OSCAR capability for debrief also appears to be suitable for extension. However, methods must be developed for presentation of multiple points of view, nuances, alternative courses of action, and prescription.

Informed instructor pilots provided subjective data that OSCAR would have utility for intercept training. It would be expected to unburden the instructor pilot and permit the AIT to be used beneficially without the immediate presence of an instructor. OSCAR received good acceptance in an environment in which there was a specific schoolhouse solution for tactical intercepts.

One may speculate that OSCAR can be used in many other training environments. For example, those in which performance issues are more vague and where users are exploring the fringes of tactical knowledge. For example, the AIT configured for basic intercepts may be extended to a device that is challenging and useful to pilots in operational squadrons. However, the question of utility, in these cases, remains to be empirically tested.

Future Directions

OSCAR may be useful for two paths of development: One avenue is to use OSCAR to develop knowledge about automated feedback and to answer some of the design decisions that had to be made during the course of this program. The other avenue is to use a variant of the existing OSCAR to test whether the OSCAR concept has utility in different and challenging training applications. For example, tasks with real-time requirements and complex maneuvering will challenge the narration capability. More difficult performance considerations and additional instructional knowledge will challenge the expert critiquing system. Any of these may challenge the presentation of feedback. In any case, there is a need to measure the quantitative effectiveness of training both with and without OSCAR, to determine the value of the concept.

REFERENCES

Carroll, J.B. (1960).Vectors of prose style. In Sebeok, T.A. (Ed.), Style in language. New York: John Wiley & Sons.

Cusella, L.P. (1984). The effects of feedback source, message and receiver characteristics on intrinsic motivation. Communication Quarterly, 32(3), 211-221.

Department of Defense (1989). Mission Employment Tactics, Fighter Fundamentals, F-16 (TAC/PACAF/USAFEM 3-3, Vol. V). May 15, 1989.

Downs, C.W., Johnson, K.M. & Barge, J.K. (1984). Communication feedback and task performance in organizations: A review of the literature. In Greenbaum, H.H., Falcione, R.L. and Hellweg, S.A. (Eds.) Organizational communication: Abstracts, analysis and overview, Vol. 9 (pp. 13-47). Beverley Hills, CA: Sage Publications.

Downs, C.W., Johnson, K.M. & Fallesen, J.J. (1987). Analysis of feedback in after action reviews (Technical Report 745). Fort Leavenworth, KS: Army Research Institute Field Unit.

Fisher, C.D. (1979). Transmission of positive and negative feedback to subordinates: A laboratory investigation. Journal of Applied Psychology, 64(5), 533-540.

Ilgen, D.R., Fisher, C.D. & Taylor M.S. (1979). Consequences of individual feedback on behavior in organizations. Journal of Applied Psychology, 64(4), 349-371.

Ilgen, D.R., Fisher, C.D. & Taylor, M.S. (1984). Performing feedback: A review of its psychological and behavioral effects (Research Note 84-47). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Jacobs, M., Jacobs, A., Feldman, G. & Cavior, N. (1973). Feedback II--The "credibility gap": Delivery of positive and negative and emotional and behavioral feedback in groups. Journal of Consulting and Clinical Psychology, 41(2), 215-223.

Miller, P.L. (1984) A critiquing approach to expert computer advice: ATTENDING. Boston: Pittman Advanced Publishing Program.

Kaplan, I.T. & Fallesen, J.J. (1986). After action review (AAR) guide for the Army Training Battle Simulation System (ARTBASS) (Research Product 86-32). Fort Leavenworth, KS: Army Research Institute Field Unit.

Nadler, D.A. (1979). The effects of feedback on task group behavior: A review of the experimental literature. Organizational Behavior and Human Performance, 22(3), 309-338.

Obermayer, R.W. & Vreuls, D. (1989). Functional description of the demonstration system for Observing System for Critique, Advice and Review (OSCAR). San Diego, CA: VRC Corporation.

Osgood, C.E., Suci, G.J., & Tannenbaum, P.H. (1957). The measurement of meaning. Urbana, IL: University of Illinois Press.

Scott, T.D. & Fobes, J.L. (1982). After action review guidebook I: National Training Center (Research Product 83-11). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

VRC Corporation (1990a). Description of Observing System for Critique, Advice and Review (OSCAR): Air Intercept Trainer (AIT) Configuration. Thousand Oaks, CA: Author.

VRC Corporation (1990b) Observing System for Critique, Advice and Review (OSCAR). Preliminary Training Test. Thousand Oaks, CA. Author.

Vreuls, D. & Obermayer, R.W. (1985). Human-System Performance Measurement in Training Simulators. Human Factors, 27(3), 241-250.

APPENDIX A

OSCAR GUIDED TOUR

Enter Initials

Please use the keyboard to enter three letters of the alphabet corresponding to your initials. Please use lower case letters only. AIT engagements will be added to OSCAR in sequence, identified with your initials and a sequence number; e.g., aaa1, aaa2, aaa3, ...

Use of the Mouse

All other control of OSCAR will be done with the mouse. The mouse is used to move the arrow on the screen as a way to point, while the mouse buttons are used to select an option for execution. Note that the mouse has three buttons: LEFT/MIDDLE/RIGHT. The mouse must be used on the special pad.

Display Buttons

There are two small areas on the initial screen labeled ADD ENGAGEMENT and DEBRIEF. These are examples of display "buttons" that are used to control OSCAR. When a display button is selected using the mouse, the color and/or shading will change to indicate that it has been activated.

ADD ENGAGEMENT is used to add an intercept engagement from the AIT into OSCAR. A 5-1/4" floppy disk must be taken from the AIT and inserted into the OSCAR disk drive before this is selected. This option will be discussed further later in this material. HOWEVER, FOR THE CURRENT OSCAR ASSESSMENT, THIS WILL BE DONE FOR YOU.

USE THE MOUSE TO POINT TO THE DEBRIEF DISPLAY BUTTON, AND CLICK THE LEFT MOUSE BUTTON.

Menus

A menu of exercises will appear when engagements have been added. The mouse is used to point to the engagement you wish to debrief. At this time there is only one file, labelled DEMO.

MOVE THE MOUSE UNTIL THE NAME DEMO IS HIGHLIGHTED. CLICK THE LEFT MOUSE BUTTON TO DEBRIEF THE DEMO ENGAGEMENT.

Top Line of Debrief Display

10NM/20NM/40NM/80NM/160NM. The range scale of the B-scope radar display, just below these display buttons, is automatically scaled to correspond to that which existed during the engagement. However, the 10NM-160NM display buttons can be used to override the automatic settings as you wish during the debrief.

PAUSE/SLOW/NORMAL/FAST. OSCAR can be used to play back an engagement with these controls used like a conventional tape recorder. NOTE: There are other control options for skipping from place to place in the engagement; these will be discussed later.

NEW/RESET/QUIT. The NEW display button will bring up a menu of other exercises performed by you. Selecting one of the other exercises will result in the present playback being terminated, and the new exercise started. RESET will restart the exercise at the beginning at any time. QUIT is used to stop when you wish to fly more engagements on the AIT, or when you are finished with all debriefing.

EXPERIMENT WITH THE PAUSE, SLOW, NORMAL, FAST, AND RESET DISPLAY BUTTONS AT THIS TIME.

Intercept Playback Displays

The top displays present information about the intercept while OSCAR is playing back recorded engagements. These displays should be familiar to you as they are very similar to those on the AIT. Some differences should be noted. Targets, radar modes, and min/max beam altitudes are displayed at all times regardless of beam orientation, lock-on, etc. The *Scores* listed with ordinary aircraft variables are rule-of-thumb estimates of probability of hit when a weapon is fired; more will be presented on this later.

Time-Line Control at the Bottom of the Screen

The time-line display at the bottom of the screen behaves like a horizontal thermometer, and shows how much playback has occurred and how much remains. However, *the timeline can* also be used to control OSCAR playback. If you point to a place on the timeline and click the left mouse button, OSCAR will re-position the playback to this point.

USE THE MOUSE TO POINT TO A PLACE NEAR THE HALFWAY POINT OF THE TIMELINE, CLICK THE LEFT MOUSE BUTTON
AFTER OSCAR REPOSITIONS, POINT TO A PLACE A LITTLE FARTHER INTO THE ENGAGEMENT, CLICK THE LEFT MOUSE BUTTON
AFTER OSCAR REPOSITIONS, POINT TO A PLACE EARLIER IN THE ENGAGEMENT, CLICK THE LEFT MOUSE BUTTON
Note, in this case, OSCAR resets to the beginning before repositioning.

There are display buttons next to the timeline corresponding to ranges from the initial target, and also at the points where shots were taken.
USE THE MOUSE TO POINT TO ONE OF THE RANGE OR SHOT DISPLAY BUTTONS NEXT TO THE TIMELINE, AND CLICK THE LEFT MOUSE BUTTON.

Critique Displays

A primary feature of OSCAR is its ability to comment about performance. Brief critique statements appear at the left-hand, lower display window.

Each critique has a display button next to it. Clicking on one of the critique display buttons produces a menu with three alternatives: FURTHER EXPLANATION, RULE, and GO TO TIME.

Clicking the FURTHER EXPLANATION option will result in text being displayed in the next window which provides related information from TAC/PACAF/USAFEM 3-3, Vol. V.

CLICK THE MOUSE BUTTON ON "FURTHER EXPLANATION" IN THE MENU.

CLICK THE TIMELINE NEAR THE END OF THE TIMELINE, IF NECESSARY, TO PRODUCE A CRITIQUE MESSAGE FOR A SHOT. CLICK THE DISPLAY BUTTON NEXT TO THE SHOT CRITIQUE, AND SELECT FURTHER EXPLANATION (THIS WILL GIVE INFORMATION ON THE SCORE AND DELAY MENTIONED IN THE CRITIQUE MESSAGE).

Note: The shot score and delay are not enabled.

Clicking the RULE option will present the machine rules which result in the critique message; this will appear in the far right window.

Click the Mouse Button to Produce Rules

Clicking the GO TO TIME option will result in OSCAR resetting the playback to the time at which the critique message was displayed.

CLICK THE MOUSE BUTTON TO GO TO TIME OF THE CRITIQUE.

Text Window Control

Sometimes there is too much text to be displayed in the window area provided for it and some scrolls out of view before it can be read. Note that each of the text windows has a thermometer-like area at the left edge of the text window. This is used to move text up and down in the window.

EXPERIMENT BY USING THE MOUSE TO POINT AT THE THERMOMETER AT THE LEFT EDGE OF A TEXT WINDOW AND CLICK THE MIDDLE MOUSE BUTTON.

Note: The shaded area represents the portion of the text file which is being shown in the window.

Strategies for Using OSCAR

Frankly, we do not know the best way to use OSCAR at this time. We hope you can help us in this regard. At this time, please consider how you might use OSCAR in performing a debrief, and try out some tentative strategies. We assume you will develop a procedure later during the debriefing process.

For your information, consider the following alternatives:

- Let OSCAR play in normal, or fast, mode until near the end; then investigate portions of the intercept which are interesting.
- Click the timeline near the end, read all of the critique messages at once. Use the GO TO TIME option for selected messages to investigate them more fully.
- Click the timeline at key points, use the SLOW mode or click in small increments along the timeline.

PLEASE TRY SOME CANDIDATE STRATEGIES USING THE DEMONSTRATION ENGAGEMENT, THEN CLICK THE QUIT DISPLAY BUTTON.

THANK YOU !

APPENDIX B

ASSESSMENT MATERIALS

Introduction

You are being asked to use and assess a device called OSCAR (Observing System for Critique, Advice and Review). OSCAR uses modern computing techniques to present feedback on performance during training. OSCAR is an initial, one-of-a-kind research and development product; this is the first limited test. After a brief guided tour of OSCAR's features, you will be asked to fly some air intercepts and debrief yourself using OSCAR. After that, you will be asked to fill out a questionnaire and give your assessment.

Basically, we wish to determine: (1) is the concept of real value for this type of training, and, if so, (2) how could OSCAR be improved? The focus, at this time, is on general characteristics and applicability; however, we will also be grateful for comments with regard to specific omissions and errors.

The air intercepts will be flown on the AIT in pairs, then OSCAR will be used to debrief the pair of intercepts. All intercepts will be 1v2 and will be flown under the assumption of no GCI and low-threat, low-technology bandits. If you have instructor-level experience in performing tactical intercepts, please fly pairs of identical intercepts and try to make one intercept of the pair "good" and the other "bad" (that is, with intentional common errors).

The Guided Tour
as presented in Appendix A
was included in the package
at this point

EXPERIENCE

INITIALS:

RANK:

TIME IN SERVICE:

F-16 HOURS:

AIR INTERCEPT TRAINER HOURS:

INSTRUCTOR-LEVEL INTERCEPT EXPERIENCE? YES__ NO__

RATING

Please circle the vertical marks (|) on the following scales to indicate your initial opinion of OSCAR.

	1	2	3	4	5	6	7	
WORTHLESS	|	|	|	|	|	|	|	VALUABLE
SHALLOW	|	|	|	|	|	|	|	DEEP
SLOW	|	|	|	|	|	|	|	FAST
VAGUE	|	|	|	|	|	|	|	PRECISE

PLEASE FLY INTERCEPTS ON THE AIT AT THIS TIME

AND DEBRIEF YOURSELF USING OSCAR

RATING

Now that you have had some hands-on experience using OSCAR, please give another set of ratings like you did before. Please circle the vertical marks (|) on the following scales to indicate your final opinion of OSCAR.

	1	2	3	4	5	6	7	
WORTHLESS	|	|	|	|	|	|	|	VALUABLE
SHALLOW	|	|	|	|	|	|	|	DEEP
SLOW	|	|	|	|	|	|	|	FAST
VAGUE	|	|	|	|	|	|	|	PRECISE

For each cell in the following FEEDBACK by ENGAGEMENT PHASE matrix, please circle your rating (1=bad, 5=good). Further, please comment on your rating, and specific good/bad features. Continue on the back side of this sheet if you need more room.

FEEDBACK CATEGORY	SORT/TARGET	INTERCEPT	WPN DELIV.
INFOR-MATION CONTENT	BAD-1-2-3-4-5-GOOD COMMENTS?:	BAD-1-2-3-4-5-GOOD COMMENTS?:	BAD-1-2-3-4-5-GOOD COMMENTS?:
PRESEN-TATION OF INFOR-MATION	BAD-1-2-3-4-5-GOOD COMMENTS?:	BAD-1-2-3-4-5-GOOD COMMENTS?:	BAD-1-2-3-4-5-GOOD COMMENTS?:
CAPABILITY TO EXPLORE & ANALYZE	BAD-1-2-3-4-5-GOOD COMMENTS?:	BAD-1-2-3-4-5-GOOD COMMENTS?:	BAD-1-2-3-4-5-GOOD COMMENTS?:

For each of the following DESIGN FEATURES, please circle your rating (1=bad, 5=good). Further, please comment on your rating, and specific good/bad features. Continue on the back side of this sheet if you need more room.

OSCAR FEATURE	RATING AND COMMENTS
INITIAL LOG-ON SCREEN	BAD-1-2-3-4-5-GOOD COMMENTS?:
TOP: PLAYBACK CONTROL BUTTONS	BAD-1-2-3-4-5-GOOD COMMENTS?:
INTERCEPT PLAYBACK DISPLAYS	BAD-1-2-3-4-5-GOOD COMMENTS?:
CRITIQUE MESSAGES, EXPLANA-TIONS AND RULES	BAD-1-2-3-4-5-GOOD COMMENTS?:
BOTTOM: TIME-LINE CONTROL	BAD-1-2-3-4-5-GOOD COMMENTS?:

1. What procedure did you develop for using OSCAR?

2. (a) Would you think that there would be good or bad training effects from using OSCAR?

   (b) How do you think training would be impacted?

   (c) For what kind/level of training would OSCAR be best (worst)?

   (d) Would OSCAR be better for individual self-debrief, or for instructor-student debrief?

3. (a) What is the best feature of OSCAR?

   (b) What is the worst feature of OSCAR?

   (c) Specific errors or bugs?

4. Any further comments? Anything we did not cover?

<u>**THANK YOU**</u>